DOCUMENT RESUME

ED 149 621 FL 009 208

AUTHOR Bikson, Tora K.
TITLE Do They Talk the Same Language? Lexical Interface and Ethnicity.
INSTITUTION Rand Corp., Santa Monica, Calif.
PUB DATE Aug 77
NOTE 21p.
AVAILABLE FROM Publications Department, The Rand Corporation, 1700 Main Street, Santa Monica, California 90406 ($1.50)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Black Youth; Caucasian Students; *Child Language; Elementary Education; Elementary School Students; *Ethnic Groups; Ethnicity; Language Research; *Language Styles; *Language Usage; Lexicology; Mexican Americans; Minority Groups; Spanish Speaking; Speech Communication; *Vocabulary

ABSTRACT

Spontaneous speech performance of ethnically diverse subjects was investigated by measuring within- and between-group lexical interface and by assessing direction of vocabulary overlap. Interview data were collected from 144 elementary school children, comprising equal white, chicano and black subsamples evenly divided among lower and upper grades. Analyses focus on whether texts exhibit a single lexicon better accessed by white subjects or, instead, generate plural lexica accessed with comparable adequacy by each ethnic group. Similar within-group lexical styles emerged. However, interface exploration found minority lexica to be different from and broader than the white lexicon they partially subsume. (Author/AM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED149621

FL

DO THEY TALK THE SAME LANGUAGE? LEXICAL INTERFACE AND ETHNICITY

Tora K. Bikson

August 1977

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
The Rand Corporation
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

FL009208

P-5996
$1.50

The Rand Paper Series

Papers are issued by The Rand Corporation as a service to its professional staff. Their purpose is to facilitate the exchange of ideas among those who share the author's research interests; Papers are not reports prepared in fulfillment of Rand's contracts or grants. Views expressed in a Paper are the author's own, and are not necessarily shared by Rand or its research sponsors.

The Rand Corporation
Santa Monica, California 90406

DO THEY TALK THE SAME LANGUAGE? LEXICAL INTERFACE AND ETHNICITY

Spontaneous speech performance of ethnically diverse subjects is investigated by measuring within- and between-group lexical interface and by assessing direction of vocabulary overlap. Interview data were collected from 144 elementary school children, comprising equal white, chicano and black subsamples evenly divided among lower and upper grades. Analyses focus on whether texts exhibit a single lexicon better accessed by white Ss or rather generate plural lexica accessed with comparable adequacy by each ethnic group. Similar within-group lexical styles emerged. However, interface exploration found minority lexica to be different from and broader than the white lexicon they partially subsume.

Tora Kay Bikson
The Rand Corporation

DO THEY TALK THE SAME LANGUAGE? LEXICAL INTERFACE AND ETHNICITY

Tora Kay Bikson
The Rand Corporation

Lexical style in the spontaneous speech of ethnically diverse elementary school students merits investigation for at least two reasons. First, research related to standardized vocabulary testing consistently reports significant race- and class-differentiated outcomes across a wide variety of instruments favoring white higher-status students (Severson and Guest, 1970). For minority students these results have distressing consequences in relation to achievement (e.g., Chall (1958) has established that reading comprehension is largely a function of vocabulary) and to inferred ability (e.g., DiLorenzo and Brady (1969) have demonstrated that vocabulary accounts for a considerable proportion of the variance in intelligence test scores). Second, although elementary grades span the period when most lexical development occurs, status differentiation on vocabulary measures becomes more pronounced as grade level increases (Deutsch, 1965). Developmental trends do nothing to close the lexical gap between minority and white students.

Explanations for such results commonly presuppose that lower SES and minority students come to school with "less language" (Raph, 1965), that they lack middle class children's "skill of choosing a best-fit response" (John, 1963). Increasing discrepancy with grade level is then accounted for in terms of "cumulative" deficit (Deutsch, 1965) or delayed "linguistic maturity" (Entwisle, 1968). An alternative hypothesis is that, in sampling the semantic systems of minority or lower

status children in comparison with white or higher status children, different working lexicons are being sampled. While there is no body of evidence directly related to this question, the concept of different partially-overlapping lexica is plausible. First, similarly pluralistic hypotheses have been confirmed for phonological and syntactic systems. Second, it accounts for grade level discrepancies in the following way. Less status differentiation is found among first-graders because at that age response variability is greater for all speakers. As speakers get older, responses standardize (or "mature") in the direction of adult responses. Given the current hypothesis, responses would standardize in different directions approximating different lexicons; and since higher status children approximate the preferred English lexicon (the lexicon over which tests are normed), measurable superiority of higher status speakers would increase. Scores so derived would not adequately access minority lexical style; however, in spontaneous speech, ethnically diverse students would be expected to reveal comparably extensive, although different, working vocabularies.

Method

Subjects. Ss were 72 lower-grade (third-fourth) and 72 upper-grade (fifth-sixth) elementary school children. Grade levels included equal numbers of white, chicano and black Ss, sex balanced within each ethnic group.[1]

---

[1]Data were collected as part of a larger study of school desegregation, NSF Grant 4-444040-22064 (H. B. Gerard and N. Miller, principal investigators).

Procedure. In an interview situation, each S told stories to each of four cartoon-like pictures from the Children's Apperception Test. Card sequence was standardized, and Ss were encouraged to continue until they had nothing further to say. All outputs were tape recorded.

Dependent measures. Recorded outputs were used to generate a lexicon for each S, and a lexicon for each cell (or, subject group defined by ethnicity, age and sex). By 'lexicon' is meant an alphabetically ordered list in which appears each distinct word employed (regardless of frequency of employment) by the individual or group from whose text the lexicon is compiled. These lists generated the following dependent measures. (1) Number of unique words was obtained for each S by comparing the S lexicon with relevant group lexica. Any lexical entries in an individual list which do not appear in a group list are unique relative to that group. Number of unique words was computed for each S relative to: S's proper cell lexicon (the list representing usage by all other Ss similar in ethnicity, age, and sex); and the two contrary cell lexica (lists obtained from groups of Ss similar in age and sex but differing in ethnicity). (2) Number of common words was obtained by a comparable search for intersects among individual and group lexica. For each S, number of words in common was computed relative to S's proper cell lexicon and the two contrary cell lexica. (3) Finally, lexical interface between an individual and a group was represented as the proportion of common to total entries in an S lexicon plus a relevant cell, or:

common words / (entries in S lexicon) + (entries in cell lexicon).

Such proportions provided an index of lexical interface for each S relative to S's cell lexicon as well as to each of the two contrary set lexica.

Predictions. Predictions relevant to lexical style in spontaneous speech focus on differences which would arise depending on whether ethnically diverse groups differentially access the same lexicon or whether they similarly access different lexical lists. On either view, lexical homogeneity is expected to be greater within than between groups and to increase with age, reflecting semantic maturation. However, the single lexicon hypothesis predicts more word sharing and at an earlier age among white Ss, minority Ss employing fewer words in common with same-cell Ss especially in the lower grade level. In contrast, the lexical plurality thesis predicts that, in relation to own-group lexica, minority usage is at least as homogeneous as white usage at comparable age levels. Finally, a single lexicon view would expect word-sharing to be a one-way event, with the white lexicon subsuming minority lexica. That is, if minority Ss simply possess fewer entries from the same lexicon, white Ss would employ (approximately) all words used by minority Ss and more besides; but minority Ss would employ only part of the white lexicon, and (virtually) no words outside it. Given the lexical plurality thesis, however, all group lexica are expected to interface in varying amounts, but no group lexicon is expected to comprise the comprehensive set subsuming other lexica as incomplete parts.

Results

Dependent measures were first examined in analyses of variance where ethnicity, grade level and sex served as independent factors. Of primary interest were effects of ethnicity, grade level, and their interaction. These sources of variance were expected to indicate whether minority Ss

exhibit differing degrees of lexical homogeneity and heterogenity in comparison with white age-mates over the elementary grades. Influence of sex was not explored. Planned comparisons collapsed over sex to investigate differences between groups defined by ethnicity and age. Then lexical interface scores were analyzed in a series of paired comparisons to determine whether systematic differences in direction of vocabulary overlap occurred among groups. For this purpose, each S's within-cell lexical interface index was compared to indices of interface with two contrary cell lexica using the Wilcoxen test for matched pairs of data.

Within-cell measures. Number of unique words used by an S relative to same-cell Ss was not substantially influenced by any independent factor (see Table 1). That ethnic effects did not emerge suggests that groups do not differ in lexical heterogeneity, in contrast to the hypothesis that white children converge on semantic norms while minority children's language use remains diffuse. A complementary picture is provided by examining number of common words, which indicates extent of word-sharing within cell (see Table 2). On this measure an ethnic effect did appear ($p < .05$), establishing that black and chicano Ss actually employed a greater number of lexical entries in common than the white sample. In addition, an approximately significant race-by-grade interaction term ($.05 < p < .10$) located age effects in the white sample only; black and chicano Ss exhibited greater lexical homogeneity than white Ss at the younger grade level, this difference decreasing as the white lexica became more homogeneous with age. Finally, the lexical interface index was also examined because, in employing the proportion of common to total lexical entries, it eliminates differences in scores reflective of

Table 1

Number of Unique Words Relative to
Same Cell Lexica

	White $\bar{X}$ = 29.6		Chicano $\bar{X}$ = 34		Black $\bar{X}$ = 33.9	
	male	female	male	female	male	female
young:	$\bar{X}$ = 26	$\bar{X}$ = 22.7	$\bar{X}$ = 33	$\bar{X}$ = 35.1	$\bar{X}$ = 30.5	$\bar{X}$ = 34.2
X=30.2	sd= 18.6	sd= 12.7	sd= 24.9	sd= 22.1	sd= 20.9	sd= 39.9
old:	$\bar{X}$ = 32.7	$\bar{X}$ = 36.9	$\bar{X}$ = 32.6	$\bar{X}$ = 35.2	$\bar{X}$ = 31.9	$\bar{X}$ = 39
$\bar{X}$=34.7	sd= 27.8	sd= 42.8	sd= 30.6	sd= 44.9	sd= 24.7	sd= 39.6

male $\bar{X}$ = 31.1 female $\bar{X}$ = 33.9 grand $\bar{X}$ = 32.5

Table 2

Number of Common Words Relative to
Same Cell Lexica

	White $\bar{X}$ = 109		Chicano $\bar{X}$ = 131		Black $\bar{X}$ = 131	
	male	female	male	female	male	female
young: $\bar{X}$=119	$\bar{X}$ = 91 sd= 27	$\bar{X}$ = 95 sd= 27	$\bar{X}$ = 120 sd= 44	$\bar{X}$ = 156 sd= 46	$\bar{X}$ = 121 sd= 49	$\bar{X}$ = 129 sd= 56
old: $\bar{X}$=128	$\bar{X}$ = 122 sd= 31	$\bar{X}$ = 126 sd= 51	$\bar{X}$ = 130 sd= 50	$\bar{X}$ = 116 sd= 47	$\bar{X}$ = 124 sd= 38	$\bar{X}$ = 150 sd= 74

male $\bar{X}$ = 118 female $\bar{X}$ = 129 grand $\bar{X}$ = 123

Ethnicity: F = 3.24
Grade Level: F = 1.32
Sex: F = 1.7
Ethnicity X Grade Level: F = 2.74
Ethnicity X Sex: F = .21
Grade Level X Sex: F = .41

absolute difference in response length (minority Ss tend to employ more words in the spontaneous speech task [Bikson, 1974]). On this measure, no independent sources of variance proved significant (see Table 3). It appears, then, that minority Ss either converge on common group lexica earlier than white children, or else that this phenomenon is an artifact of response length such that diverse ethnic groups manifest similar within-group patterns of linguistic homogeneity and heterogeneity.

Between-cell measures. A second set of analyses investigated the extent to which members of one ethnic group know and use the words employed by (same-aged members of) another. Number of words used uniquely by an S relative to contrary groups was susceptible to ethnic influence ($p \simeq .10$), an effect attributable primarily to between-group differences at the younger grade level; these results indicated that young minority Ss use more words not employed outside their own ethnic groups than do their white counterparts (see Table 4). On the other hand, examining the number of words used in common by an S and members of other ethnic groups revealed that young minority speakers employ a greater number of items jointly with other ethnic groups as well (see Table 5). Finally, the between-cell lexical interface index also showed ethnic effects ($.05 < p < .10$) favoring minority Ss (see Table 6). That is, controlling for response length, black and chicano Ss were nevertheless found to use significantly more words from other group lexica than did white Ss. A race-by-grade interaction ($.05 < p < .10$) again located the ethnic chiefly at the lower age level.

Direction of lexical overlap. Paired comparisons established that only the white sample evidences the predicted greater lexical overlap

Table 3

Lexical Interface, or Percent of Common Words
Relative to Same Cell Lexica

	White $\bar{X}$ = 13.08		Chicano $\bar{X}$ = 13.33		Black $\bar{X}$ = 13.23	
	male	female	male	female	male	female
young:	$\bar{X}$ = 12.92	$\bar{X}$ = 13.58	$\bar{X}$ = 13.0	$\bar{X}$ = 14.75	X = 13.5	X = 13.08
$\bar{X}$=13.47	sd= 3.2	sd= 3.5	sd= 3.98	sd= 3.81	sd= 4.84	sd= 4.77
old:	$\bar{X}$ = 13.5	$\bar{X}$ = 12.33	$\bar{X}$ = 13.42	$\bar{X}$ = 12.17	$\bar{X}$ = 13.33	$\bar{X}$ = 13.0
$\bar{X}$=12.96	sd= 2.63	sd= 3.88	sd= 4.17	sd= 3.83	sd= 3.17	sd= 5.26

male $\bar{X}$ = 13.28 female $\bar{X}$ = 13.15 grand $\bar{X}$ = 13.21

Table 4

Average Number of Unique Words Relative to
The Two Contrary Cell Lexica

	White $\bar{X}$ = 24.08		Chicano $\bar{X}$ = 35.12		Black $\bar{X}$ = 35.85	
	male	female	male	female	male	female
young:	$\bar{X}$ = 18.58	$\bar{X}$ = 15.58	$\bar{X}$ = 35.92	$\bar{X}$ = 43.92	$\bar{X}$ = 32.42	$\bar{X}$ = 36.33
$\bar{X}$=30.46	sd= 14.84	sd= 8.14	sd= 25.63	sd= 24.21	sd= 23.64	sd= 38.74
old:	$\bar{X}$ = 27.83	$\bar{X}$ = 34.33	$\bar{X}$ = 31.17	$\bar{X}$ = 29.5	$\bar{X}$ = 30.0[illegible]	$\bar{X}$ = 44.58
$\bar{X}$=32.92	sd= 22.93	sd= 36.62	sd= 29.36	sd= 36.38	sd= 22.[illegible]1	sd= 41.64

male $\bar{X}$ = 29.33 female $\bar{X}$ = 34.04 grand $\bar{X}$ = 31.64

Ethnicity: F = 2.32
Grade Level: F = .24
Sex: F = .88
Ethnicity X Grade Level: F = 1.85
Ethnicity X Sex: F = .20
Grade Level X Sex: F = .12

Table 5

Average Number of Common Words Relative to
The Two Contrary Cell Lexica

	White $\bar{X}$ = 113.75		Chicano $\bar{X}$ = 128.94		Black $\bar{X}$ = 124.44	
	male	female	male	female	male	female
young:	$\bar{X}$ = 97.64	$\bar{X}$ = 101.92	$\bar{X}$ = 116.75	$\bar{X}$ = 146.67	$\bar{X}$ = 118.33	$\bar{X}$ = 126.33
$\bar{X}$=117.94	sd= 31.42	sd= 31.71	sd= 42.77	sd= 43.45	sd= 45.93	sd= 55.88
old:	$\bar{X}$ = 127	$\bar{X}$ = 128.42	$\bar{X}$ = 130.67	$\bar{X}$ = 121.67	$\bar{X}$ = 125.42	$\bar{X}$ = 143.67
$\bar{X}$=129.47	sd= 35.84	sd= 56.63	sd= 50.57	sd= 51.1	sd= 39	sd= 71.67

male $\bar{X}$ = 119.3 female $\bar{X}$ = 128.1 grand $\bar{X}$ = 123.71

Ethnicity: F = 2.64
Grade Level: F = 1.36
Sex: F = .54
Ethnicity X Grade Level: F = 1.21
Ethnicity X Sex: F = .58
Grade Level X Sex: F = .14

Table 6

Lexical Interface, or
Average Percent of Common Words, Relative to the Two
Contrary All Lexica

	White $\bar{X}$ = 11.71		Chicano $\bar{X}$ = 13.98		Black $\bar{X}$ = 13.875	
	male	female	male	female	male	female
young:	$\bar{X}$ = 11	$\bar{X}$ = 10.42	$\bar{X}$ = 14.08	$\bar{X}$ = 16.75	$\bar{X}$ = 14	$\bar{X}$ = 14.33
$\bar{X}$=13.43	sd= 3.16	sd= 2.9	sd= 4.25	sd= 4.15	sd= 4.58	sd= 4.95
old:	$\bar{X}$ = 13.42	$\bar{X}$ = 12	$\bar{X}$ = 13.92	$\bar{X}$ = 11.17	$\bar{X}$ = 13.33	$\bar{X}$ = 13.83
$\bar{X}$=12.94	sd= 2.96	sd= 4.06	sd= 4.07	sd= 4.28	sd= 3.37	sd= 5.55

male $\bar{X}$ = 13.29 female $\bar{X}$ = 13.08 grand $\bar{X}$ = 13.19

Ethnicity: F = 4.3
Grade Level: F = .46
Sex: F = .08
Ethnicity X Grade Level: F = 3.89
Ethnicity X Sex: F = .34
Grade Level X Sex: F = 2.01

within than between cells ($p < .05$) (see Table 7). That is, especially at younger grade levels, a white S's lexicon is much more cohesive with the lexica of white age-mates than with vocabularies of minority Ss. In the black and chicano samples, on the other hand, more significant comparisons indicated greater overlap with a contrary cell lexicon than with S's own cell lexicon. (Previous analyses preclude attributing such results to ethnic differences in within-group lexical homogeneity and heterogeneity.) Chicano lexica tended to interface with both white and black vocabularies, while black Ss' vocabularies more often overlapped with white usage. For minority speakers generally, interface with the white vocabulary may well represent a relatively successful attempt to incorporate the socially preferred lexicon in a school-like spontaneous speech task.

Discussion

Analyses of variance in conjunction with paired comparisons of lexical interface measures indicated that in fact ethnically diverse speakers make different kinds of lexical choices, particularly in the younger grades. While all groups exhibited similar within-cell lexical homogeneity, ethnic differences emerged when extent of interface with contrary cell lexica was considered. These data disconfirmed predictions from a single lexicon thesis and corroborated a pluralistic view: while the white lexicon did not stabilize earlier and subsume minority lexica, the converse often held. Further, while black and chicano lexica differed, these differences seemed not as extensive as those between the white lexicon and either minority lexicon. In any case, this research

Table 7

Wilcoxen T Values for Same - vs. Contrary-Cell Lexical Interface Comparison [1]

Cells (n = 12)	T
Young White Male Lexica	
1) Within vs. young chicano male	0**
2) Within vs. young black male	0**
Old White Male Lexica	
1) Within vs. old chicano male	31.5
2) Within vs. old black male	37.5
Young White Female Lexica	
1) Within vs. young chicano female	1.5**
2) Within vs. young black female	1.5**
Old White Female Lexica	
1) Within vs. old chicano female	31
2) Within vs. old black female	21
Young Chicano Male Lexica	
1) Within vs. young white male	8*
2) Within vs. young	19.5
Old Chicano Male Lexica	
1) Within vs. old white male	5**
2) Within vs. old black male	10.5**
Young Chicano Female Lexica	
1) Within vs. young white female	0**
2) Within vs. young black female	6**

Table 7 (cont'd)

Cells (n = 12)	T
Old Chicano Female Lexica	
1) Within vs. old white female	13*
2) Within vs. old black female	17.5[t]
Young Black Male Lexica	
1) Within vs. young white male	3.5**
2) Within vs. young chicano male	12*
Old Black Male Lexica	
1) Within vs. old white male	20
2) Within vs. old chicano male	31
Young Black Female Lexica	
1) Within vs. young white female	0**
2) Within vs. young chicano female	22.5
Old Black Female Lexica	
1) Within vs. old white female	0**
2) Within vs. old chicano female	6.5**

[1]A subject's within-cell lexical interface score is contrasted with his interface score in relation to each appropriate contrary cell lexicon. This procedure yields two pairs of matched scores for each subject in a cell. For example, within-cell lexical interface measures on each young white male are compared first with his lexical interface scores relative to the young chicano male lexicon and are then compared with his vocabulary interface scores relative to the young black male lexicon; and a similar group of comparisons is made for each cell, member by member, in relation to its appropriate contrary cells. What results are twelve sets of paired comparisons, and for each comparison n = 12. The Wilcoxen Test for matched pairs of data (Siegel, 1956) is used to determine the extent to which subject lexicons systematically differ from those of contrary cells. Significant values of T (Wilcoxen) indicate systematic differences in direction of lexical overlap; nonsignificant values show that extent of overlap does not differ regardless of whether an S's lexicon is compared with that of his own or an ethnically distant contrary cell. The table presents T values for all such comparisons in twelve sets, where ** = $p < .01$; * = $p < .02$; and t = $.05 < p < .10$. A two-tailed test is used because, while it had been expected that within-cell overlap would always be greater than cross-cell overlap, prior analyses of variance made this general prediction questionable. It could be expected to hold for white groups but not necessarily for minority groups, which exhibited significantly higher contrary cell lexical interface scores. Underlined significant T values represent cases where the within-cell interface score is greater than the contrary cell interface score; otherwise, significant values indicate overlap with the contrary cell is greater than overlap with the subject's own cell lexicon.

supports the thesis that speakers from diverse ethnic groups access different but partially overlapping lexica. While the language of black and chicano Ss in spontaneous speech reveals a broader working vocabulary than that of white age-mates, assessments based on majority usage maintain race and class differentiation of outcomes by failing to represent the range of minority lexical style.

REFERENCES

Bikson, T. K. Minority speech as objectively measured and subjectively evaluated. Personality and Social Psychology Bulletin, 1974, 1, 336-338.

Chall, J. S. Readability: An appraisal of research and application. Columbus, Ohio: Bureau of Education Research, Ohio State University, 1958.

Deutsch, M. P. The role of social class in language development and cognition. American Journal of Orthopsychiatry, 1965, 25, 78-88.

Dilorenzo, L. T. and Brady, J. J. Use of the Peabody Picture Vocabulary Test with preschool children, Training School Bulletin, 1969, 117-121.

Entwisle, D. Subcultural differences in children's language development. Internal Journal of Psychology, 1968, 3, 13-22.

John, V. The intellectual development of slum children: Some preliminary findings. American Journal of Orthopsychiatry, 1963, 33, 813-822.

Ralph, J. Language development in socially disadvantaged children. Review of Educational Research, 1965, 32, 203-214.

Severson, R. and Guest, K. Toward the standardized assessment of the language of disadvantaged children. In F. William (Ed.), Language and Poverty. Chicago: Markham Press, 1970.